

Abstrações – modelo de dados - projeto conceitual e lógico.

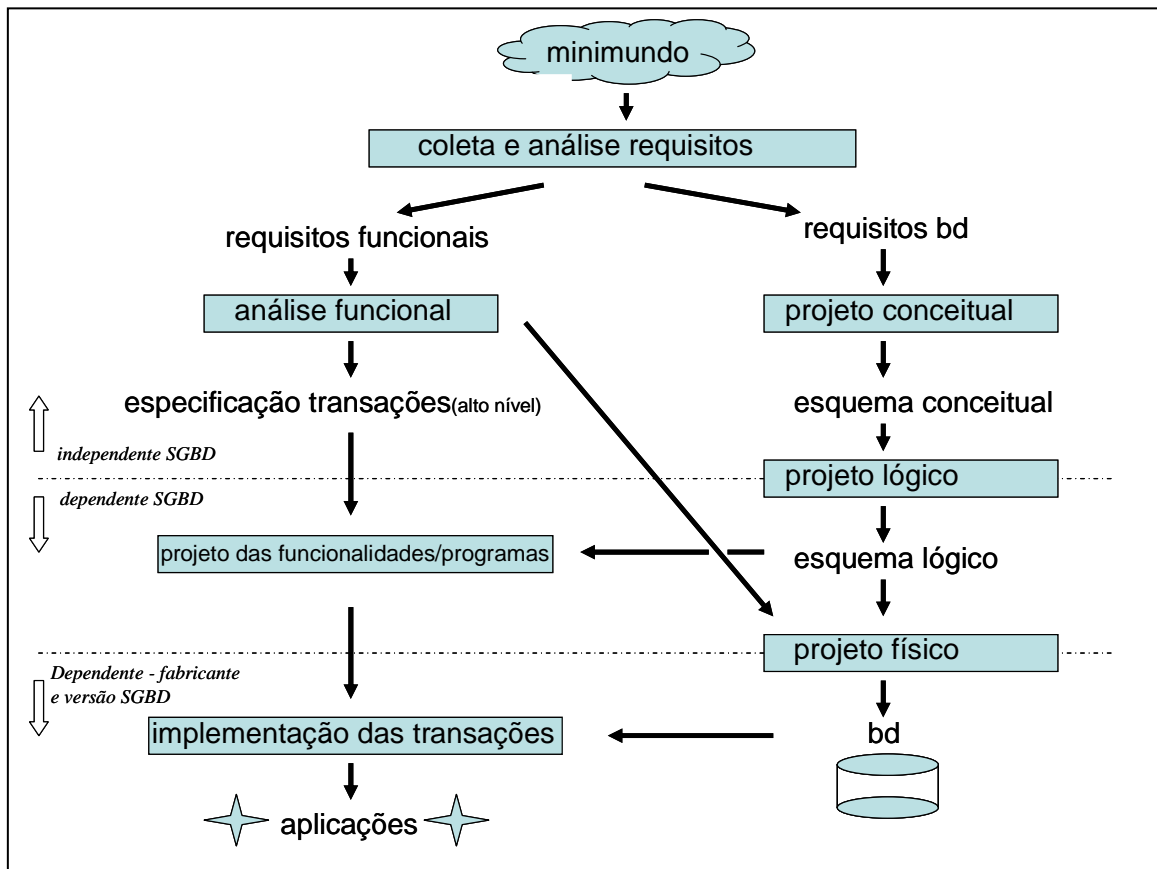


Fig1: diagrama simplificado para ilustrar as principais fases do projeto BD.

Definição de conceitos.

Banco de dados (BD) → conjunto de dados integrados que por objetivo atender a uma comunidade de usuários.

Modelo de dados → descrição formal das estruturas de dados para representação de um banco de dados; com suas respectivas restrições e linguagem para criação e manipulação de dados. Denominaremos de construtores semânticos à descrição formal para cada estrutura representada no modelo de dados. Exemplos de modelos de dados: Relacional, Orientado a Objetos, Rede, Hierárquico, Entidade-Relacionamento.

Sistema de gerência de banco de dados (SGBD) → software que incorpora as funções de definição, recuperação e alteração de dados em um banco de dados. Os SGBDs implementam os modelos de dados.

Modelagem de dados → é a ação de representar/abstrair os objetos de dados do minimundo com o objetivo da criação dos projetos conceitual e lógico de um BD. Alguns autores incluem o Projeto físico como parte da modelagem de dados pelo fato de que as otimizações são oriundas de análises do comportamento dinâmico do banco de dados.

Projeto conceitual BD → esquema de dados abstratos que descreve a estrutura de um banco de dados de forma independente de um SGBD.

Projeto lógico BD → esquema de dados que representa a estrutura de dados de um banco de dados em acordo com o modelo de dados subjacente a um SGBD.

Projeto físico BD → esquema de dados do projeto lógico adicionado às estratégias de otimização para manipulação das estruturas de dados. As estratégias de otimização são dependentes dos fabricantes dos SGBDs e de suas versões.

As abstrações de dados (inerentes aos construtores semânticos dos modelos de dados) são as formas de representar os conceitos subjacentes aos de requisitos de projeto de banco de dados. Tais requisitos são extraídos do minimundo. Assim para a modelagem de dados (projeto conceitual e lógico) é fundamental o domínio dos construtores semânticos dos modelos de dados. Um modelo de dados que merece destaque é o Entidade-Relacionamento.

Entidade → conjunto de objetos da realidade modelada, sobre os quais se deseja manter informações no banco de dados.

Relacionamento → conjunto de associações entre entidades

Atributo → dado que é associado a cada ocorrência (instância) de uma entidade ou de um relacionamento.

Atributo identificador → Um identificador é um conjunto ou mais de atributos cujos valores servem para distinguir uma ocorrência de entidade (instância) ou relacionamento.

Atributo Compostos ou Simples (atômicos) → Atributos compostos podem ser divididos em subpartes menores, que representam outros atributos. Um exemplo de atributo composto é endereço que pode inclusive criar um hierarquia de valores compostos:

Endereço composto por: Logradouro – Cidade – Estado – Código Postal

Logradouro composto por: NomeLogradouro – NumeroLogradouro – Numero do Apto.

Atributos Multivalorados → Em alguns casos um atributo pode ter um conjunto de valores para a mesma entidade.

Conjunto de valores (domínios) dos atributos → cada atributo simples de uma entidade está associado a um conjunto de valores (ou domínio de valores), que especifica o conjunto de valores que podem ser designados para o respectivo atributo. Formalmente, um atributo A da entidade E cujo conjunto de valores é igual a V pode ser definido como uma função de E para o conjunto domínio P(V) de V:

A: E → P(V)

Esta definição cobre os atributos de valor único e multivalorado, bem como os nulos. Um valor nulo é representado por um conjunto vazio. Para atributos de valor único, A(e) tem uma restrição de valor individual para cada entidade E, enquanto que em

atributos multivalorados não existem restrições. Para um atributo composto A, o conjunto de valores V é o produto cartesiano de $P(V_1), P(V_2), \dots, P(V_n)$, onde V_1, V_2, \dots, V_n são os conjuntos de valores dos atributos componentes simples que forma A:

$$V = P(V_1) \times P(V_2) \times \dots \times P(V_n)$$

Cardinalidade (mínima, máxima) de entidade de relacionamento \rightarrow número (mínimo, máximo) de ocorrências de entidades associadas a uma ocorrência da entidade em questão através de relacionamento.

Relacionamentos ternários: Não é trivial entender o conceito de cardinalidade nos relacionamentos ternários. A estratégia para o entendimento é o uso de dependências funcionais. Exemplos que ilustram a cardinalidade: 1-1-1; 1-1-N; 1-N-P; M-N-P.

Projeto – Técnico – Notebook: 1-1-1

DFs: Id_Tecnico, Id_Projeto \rightarrow Id_Notebook

Id_Tecnico, Id_Notebook \rightarrow Id_Projeto

Id_Projeto, Id_Notebook \rightarrow Id_Tecnico

Projeto – Empregado – Localização: 1-N-1

DFs: Id_Técnico, Id_Localização \rightarrow Id_Projeto

Id_Técnico, Id_Projeto \rightarrow Id_localização

Projeto – Gerente – Engenheiro: N-1-P

DF: Id_Projeto, Id_Tecnico \rightarrow Id_Gerente

Projeto – Empregado – Perfil: M-N-P

DF: nenhuma!

Agregação (Entidade Associativa) \rightarrow Um construtor semântico que denota a existência de uma junção de elementos através de um relacionamento, e que permite que essa junção seja percebida como uma nova entidade a ser, por sua vez, relacionada à outra entidade. Um exemplo que ilustra bem a necessidade de tal construtor semântico é o relacionamento Medico x Paciente com a necessidade de controlar os medicamentos receitados na oportunidade da ocorrência do relacionamento Médico x Paciente. Não podemos vincular os medicamentos apenas ao Medico ou ao Paciente individualmente. Precisamos de uma nova entidade gerada a partir do relacionamento Médico-Paciente para vincular os Medicamentos. Assim do relacionamento Consulta é gerado a entidade Consulta, conforme ilustrado na figura abaixo, para controlar as prescrições dos medicamentos.

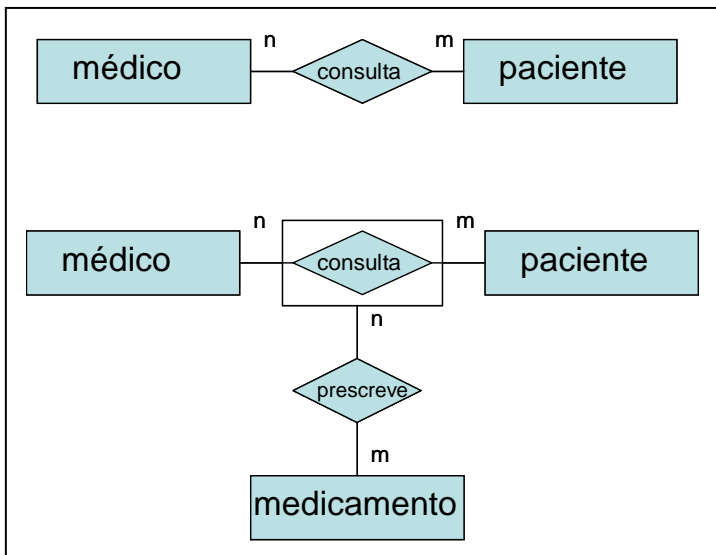


Fig2: Agregação (Entidade Associativa).

A identificação da agregação pode se tornar mais confusa quando partimos de um relacionamento triplo entre Médico – Paciente – Medicamento (com cardinalidade M-N-P). Para representar que um determinado Médico em uma determinada consulta *poderia ter prescrito* m medicamentos a um determinado paciente, seríamos obrigados a restringir a cardinalidade M-N-P (com dependências funcionais tais quais apresentadas para os relacionamentos ternários), o que de fato impõe uma grande restrição na representação semântica. Seguramente não é o caminho para solução. Com o construtor semântico agregação selecionamos um instância do relacionamento e vinculamos tal instância a uma outra entidade.

Para detalhar ainda mais o construtor semântico agregação veja os seguintes casos.

Caso 1 (A) – Agregado tem o mesmo identificador do relacionamento (ocorrência de uma única instância).

Nesse caso não é necessário criar um novo identificador para o agregado. As chaves das entidades são os identificadores. Exemplo: Professor (nroFunc, nome); Orienta(titulo); AlunoPos(nroUSP, nome, ano ingresso). Existirá um agregado denominado Orientação (título) que se apoiará no relacionamento (com identificador próprio) Professor-AlunoPos. Mesmo que seja possível identificar a agregação por um identificador próprio (atributo título), ela pode ser identificada também pelo relacionamento identificador entre as entidades que participam do relacionamento.

Caso 1 (B) – Agregado tem o mesmo identificador relacionamento (ocorrência de mais de uma instância).

Nesse caso, para identificar a entidade agregação, deve-se identificar as entidades que participam do relacionamento, além de ser necessário complementar esse identificador com atributos do relacionamento. Exemplo: Professor; Ministra; Disciplina, Agregado Aula precisa dos atributos data + hora.

Caso 2 – Agregado tem o identificador Próprio

Pode haver mais de um relacionamento com o mesmo identificador, envolvendo as mesmas entidades. Nesse caso, além de poder identificar a entidade agregação com as entidades que participam do relacionamento complementando o identificador com atributos próprios da entidade agregação, essa entidade agregada também tem um identificador próprio, “auto-suficiente”. Exemplo: Aluno; Tranca; Disciplina. O agregado trancamento tem um identificador próprio Nro_Processo.

Na figura 3 abaixo temos a ilustração desses casos. Na figura 4 temos a generalização dos casos 1, 2, 3. É importante diferenciar esses casos para o mapeamento para o modelo relacional.

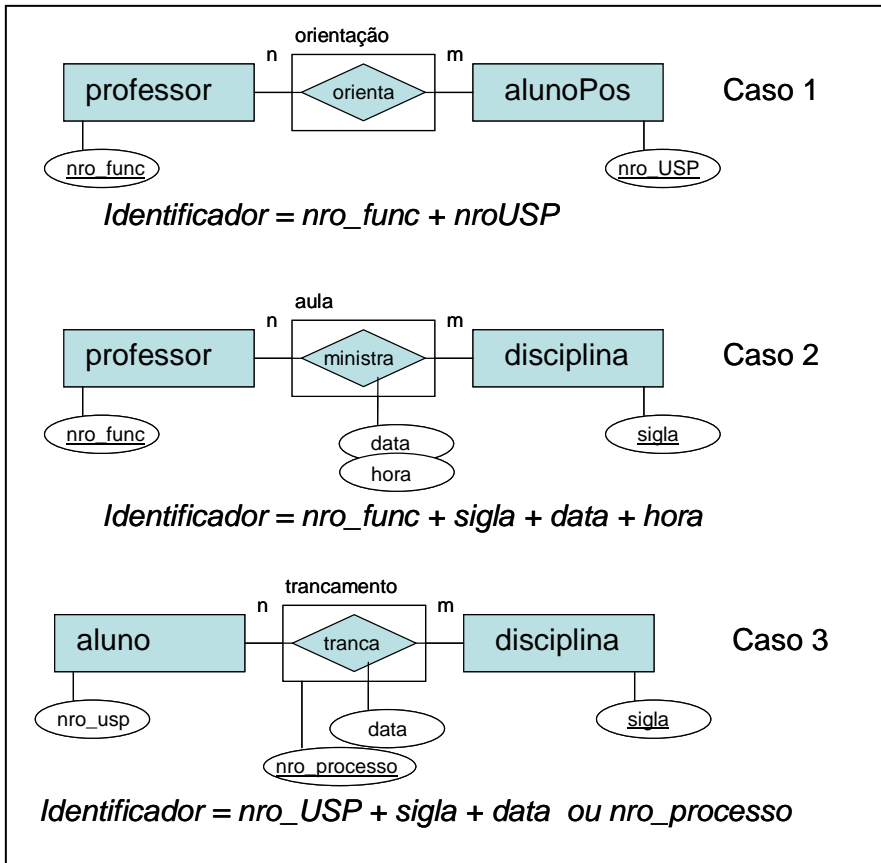


Fig 3: ilustração dos casos de agregação.

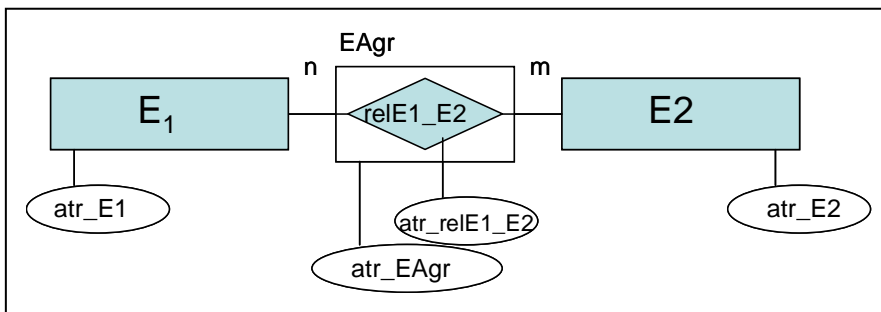


Fig 4: construtor semântico generalizando os casos 1, 2, 3.

Antes de detalhar construtores semânticos mais complexos apresentamos uma especificação textual (não completa) em BNF e também uma representação formal do modelo ER. Alguns autores separam ER e ER estendido (ERE). Nesse caso o ER teria seus construtores semânticos: Entidade, Atributo, Relacionamento e ERE: Agregação, Especialização, Categorização.

BNF (pg. 35 Heuser – Projeto de Banco de Dados)

Esquema → Esquema: Esquema_nome
SEÇÃO_ENTIDADE
SEÇÃO_AGREGAÇÃO
SEÇÃO_RELACIONAMENTO

SEÇÃO_ENTIDADE → {DECL_ENT}

DECL_ENT → Entidade: ENTIDADE_NOME
[SEÇÃO_ATRIBUTO]
[SEÇÃO_IDENTIFICADOR]

SEÇÃO_ATRIBUTO → Atributos: {DECL_ATRIB}
DECL_ATRIB → [(MIN_CARD,MAX_CARD)] ATRIBUTO_NOME [:DECL_TIPO]
MIN_CARD → 0 | 1
MAX_CARD → 1 | n
DECL_TIPO → inteiro | real | boolean | texto (INTEIRO) |
enumeração (LISTA_VALORES)

SEÇÃO_IDENTIFICADOR → Identificadores: {DECL_IDENT}
DECL_IDENT → {IDENTIFICADOR}
IDENTIFICADOR → ATRIBUTO_NOME |
ENTIDADE_NOME (via RELACIONAMENTO_NOME)

SEÇÃO_AGREGAÇÃO → {DECL_ENT_AGREGAÇÃO}
DECL_ENT_AGREGAÇÃO → EntidadeAgregação: NOME_RELACIONAMENTO

SEÇÃO_RELACIONAMENTO → {DECL_RELACION}
DECL_RELACION → Relacionamento: NOME_RELACIONAMENTO
Entidade: {DECL_ENT_RELACIONADA}
[Atributos: {DECL_ATRIB}]
[Identificadores: {DECL_IDENT}]

DECL_ENT_RELACIONADA → [(CARD_MIN,CAR_MAX)] NOME_ENTIDADE

ER formal (extrair do modelo ER e inserir neste trecho do texto)

Generalização/Especialização → Esse construtor semântico é utilizado para representar o relacionamento de uma classe e suas subclasses.

Classe $C \rightarrow C$ é um conjunto ou coleção de ocorrências (instâncias) de entidades. Essa definição é válida para qualquer um dos elementos no ER estendido que agrupam instâncias de entidades tais como: subclasses, superclasses, categorias e as próprias entidades.

Subclasse $S \rightarrow S$ é uma classe cujas instâncias das entidades devem sempre ser um subconjunto das instâncias das classes denominadas superclasse C que tem um relacionamento entre superclasse/subclasse. Esse relacionamento é representado como C/S . Para esse tipo de relacionamento C/S devemos sempre ter S contido em C .

Uma **especialização** $Z = (S_1, S_2, \dots, S_n)$ é um conjunto de subclasses que possuem a mesma superclasse G , ou seja, G/S , é um relacionamento superclasse/subclasse para $i=1,2,\dots,n$. G é chamada de classe(entidade) generalizada, ou uma generalização das subclasses $\{S_1, S_2, \dots, S_n\}$. Z é considerado total se sempre (em qualquer momento da especialização) temos:

$$US_i = G ; i=1,2,\dots,n$$

Caso contrário, Z é considerado parcial.

Z será disjuncto (exclusivo) sempre que:

S_i interseção $S_j =$ conjunto vazio para $i \neq j$ (incluir notação simbólica)

Caso contrário Z é sobreponível (tem sobreposições)

Exemplos de especializações (com variações de disjunção e sobreposição) : PESSOA
 \rightarrow Funcionário, Professor, Aluno

Uma **categoria** T é uma classe que é um subconjunto da união de n superclasses definidoras $D_1, D_2, \dots, D_n, n > 1$ e é formalmente especificada da seguinte forma:

T contido $(D_1 \cup D_2 \cup D_n)$

Uma categoria pode ser total ou parcial. Quando todas as instâncias de D_i geram uma determinada instância T temos a categoria total. Caso contrário tem-se uma categoria parcial. A palavra composição e união são usadas por alguns autores como sinônimo da palavra categoria.

Exemplos de categorização.

Total: Propriedade \rightarrow Edifício, Terreno

Bomba hidráulica \rightarrow Chassi, Motor, Acessórios

Parcial: Correntista \rightarrow Empresa, Pessoa

Cooperado \rightarrow Médico, Funcionário

Note-se que ao contrário da abstração de generalização, não existe herança de atributos na abstração de categorização (composição) e, portanto cada instância de E, componente ou composta, deve ter seu próprio identificador.

Existem muitos padrões de representação gráfica dos construtores semânticos acima apresentados, com maior ou menor grau de clareza e completude de representação dos construtores semânticos apresentados. Entretanto o mais importante é saber identificar tais construtores e homogeneizar tal representação no projeto de banco de dados. Exemplos de notações entre outros são: James Martin (Engenharia da Informação), Peter Chen, UML.

Frase para resumo de mapeamento composição:

O mapeamento de ocorrências de abstrações de composição para o Modelo Relacional tem poucas regras, uma vez que a principal motivação para essa abstração é de natureza funcional, e não estrutural.